ERDC/TEC TR-02-1

Topographic Engineering Center

US Army Corps
of Engineers®
Engineer Research and
Development Center

A Suggested Methodology for the Computation of Imagery-Based Terrain Data Reliability and Implementation Into Tactical Decision Models

Kevin R. Slocum

December 2002

A Suggested Methodology for the Computation of Imagery-Based Terrain Data Reliability and Implementation into Tactical Decision Models

By: Kevin R. Slocum, Ph.D.

Topographic Engineering Center
US Army Engineer Research and Development Center
7701 Telegraph Road
Alexandria, VA 22315-3864

Final report

Preface

This report was prepared under the Vegetation and Soil Modeling research project. This research is a consolidated collection of previously unpublished research completed between 1996 and 1999 under various work units (other than Vegetation and Soil Modeling) by Dr. Kevin R. Slocum, Vinh Duong, James Moeller, and Gery Wakefield, Information Generation and Management Branch, Topography, Imagery, and Geospatial Research Division, Topographic Engineering Center. The work was performed under the supervision of Debra Kabinier, Chief, Information Generation and Management Branch, and William Z. Clark, Jr., Chief, Topography, Imagery, and Geospatial Research Division.

Mr. Robert Burkhardt was Director and Mr. Lazlo Greczy was Deputy Director of TEC at the time of report publication.

TEC is an element of the U.S. Army Engineer Research and Development Center (ERDC), U.S. Army Corps of Engineers. The Director of ERDC is Dr. James R. Houston and the Commander is Col. John W. Morris III.

Acknowledgements

Appreciation is given to the following former TEC employees who assisted in test and evaluation of the research plans through their application of ERDAS Imagine and Arc/Info geospatial analysis software: Vinh Duong, James Moeller, and Gery Wakefield. The author also recognizes Dr. Paul Krause and Ms. Valerie Carney for their review and suggested comments for improvement of this document.

Contents

List of Figures and Tables

Figures

Tables

Topic 1. Derivation of Imagery-based Reliability Values at Terrain Theme and Pixel Levels

Note- Topic 1 has been distilled from an earlier paper presented at the August 1997 USACE Surveying Mapping and Remote Sensing Conference, St. Louis, MO (Slocum et al., 1997). Renewed interest in terrain data reliability and its impact on tactical decision aids provided compelling incentive to revisit our work from the late 1990s and repackage it into a cohesive body of research.

Background

New digital terrain data are created daily that directly support tactical decision-making. These decisions are typically not made with full understanding of the contributing terrain data quality. Data are treated as spatially invariable in quality and devoid of any metric measuring the underlying certainty of feature classification.

Digital imagery has become a preferred source from which requisite terrain features are extracted. Imagery provides a fast, nonintrusive, nonrestrictive source effectively exploited by image processing tools. Supervised classification algorithms are popular processing techniques useful for classifying an image into user-defined terrain feature classes. All picture elements (pixels) are represented by unique digital values defining the terrain conditions within that image space. Pixels are individually assigned to appropriate terrain feature classes by image processing algorithms. Commercial-off-the-shelf (COTS) image processing packages provide an opportunity to identify terrain classification reliability along with the class assignments. However, the opportunity for capturing reliability information is typically not passed along into the final terrain class map output nor is it stored as a supplementary metadata file. There does exist COTS functionality that specifically addresses image classification probability but these algorithms are dependent on *a priori* knowledge about the areas of interest to be mapped, a requirement that is often unattainable, especially overseas in denied access areas. In the absence of *a priori* information, a user may instead use the basic image processing capabilities to develop a home-grown pixel reliability method developed from a distance-to-means image processing capability. The reliability model developed for this paper focused attention on individual pixel distance-to-means values within identical terrain feature classes and on the expected separability of the various feature classes themselves.

Image sources, starting from the earliest panchromatic aerial photographs and evolving into today's sophisticated satellite imaging systems, present the image analyst with a diverse source from which geographic data may be extracted (Avery and Berlin, 1992). Satellite imagery introduced the discipline of digital image processing for geographic data classification and with it a myriad of techniques have evolved (Jensen, 1996). Identification and accurate classification of natural and cultural terrain features is an image processing goal. Image sources will continue to be a primary information source for geographic data extraction with the advent of new commercial sensor data emerging on the horizon demonstrating higher spatial resolution and continued spectral differentiation.

Potential to spectrally classify more varieties of natural and cultural features from original image source is creating profound new impacts on geographic data generation. Attempting to increase the number of terrain feature types that are classified implies a greater risk for misclassification of a feature. Capabilities are advancing quickly within the mapping discipline and with these advances come user-community expectation for accurate geographic data, or at least some measure of their reliability.

Uncertainty of geographic data is prevalent in today's age of geospatial information exchange. The degree of trust to which users associate these data varies widely from naive faith to total skepticism. There has not been a concerted effort on the part of past geographic data generators to appropriately convey the certainty of natural and cultural features classified in digital or hard-copy map space. Accuracy statements, when included with a map product, historically have taken on the form of accuracy for an entire map sheet. Variability of this accuracy within a map is not conveyed to the user.

To measure map classification accuracy, truth data of some type must exist. To measure reliability, however, there is not the same demand for rigorous truth data. Rather, reliability can be realized from statistical expectations that can be measured. A user may gain a measure of *confidence* about terrain data once provided with this information. This confidence, or reliability, can be expressed as a value that provides the datauser with information previously not included.

Academia has published fairly extensively on the subject of geographic data uncertainty, yet the implementation of these valuable ideas has not materialized in the production cycles of major geographic data producers in the private or public sectors (Strahler, 1980; Aronoff, 1982; Storey and Congalton 1986). Invariably, earlier data uncertainty work expects that ground truth is available and used in the final assessment process. With these ground truth data, consumer and producer error could be computed. *Consumer risk* is the probability that a map of unacceptable accuracy will pass an accuracy test while *producer risk* is the probability that a map of acceptable accuracy will be rejected (Aronoff, 1985). For practical purposes, collection of ground truth data can be considered impractical for areas of the world in which access is restricted or denied.

Why Consider Terrain Data Reliability?

How reliable are geospatial data that are being generated? For many data users, this question is quite important but mostly unanswered. From a military perspective, commanding officers make countless decisions that are based in large part on terrain conditions. Decision effectiveness may be vastly improved if those same commanders are provided with additional information regarding the reliability of the terrain data. Are certain parts of the map simply more reliable than others?

In the nonmilitary community, decision-making from terrain data that are devoid of reliability is equally difficult. Conclusions are drawn daily by civil and military users alike that may have serious short- and long-term implications. For example, is location A the suggested place for construction of a water runoff retention pond or is location B better? Unfortunately, decisions related to models such as site suitability, mobility, and trend analysis continue to be made without the benefit of understanding the uncertainties in the underlying data.

To illustrate the result of a short-term implication occurring because of a site suitability decision made without knowledge of spatial geographic data reliability, a military river-crossing bridge site selection is examined. A temporary bridge location is to be positioned according to "suitable" terrain conditions. All conditions are met for a dozen possible bridge crossing sites and ultimately, all conditions being equal, a location is chosen by the commanding officer that is logistically nearest to the military unit's present geographic coordinates. Information unavailable to the suitability model, and therefore the commander, is the amount of certainty that existed in the terrain conditions that ultimately guided the selection of the dozen possible locations. To improve this example, all twelve suitable site selections can be identified, followed by a prioritization of locations based on the terrain data having the highest degree of confidence. In the continuing absence of this knowledge of data

reliability, a military bridge siting or similar decision may occur at what could be the least desirable of the possible locations.

Objectives

The ultimate project goal was to develop a measure of reliability that described the confidence of terrain classes derived from imagery. Ground truth data were purposefully *not* used in development of a repeatable method of measuring reliability. Rather, a methodology was developed to derive terrain feature class reliability and subsequent within-class pixel reliability that utilized only the image data available. This does not suggest that ground truth data are without value. Undoubtedly, ground truth data should improve feature classification but the realization for users in the Armed Forces is that these truth data may not be available, yet some measure of data reliability is still demanded for informed decision making. In this project, a model to measure terrain data reliability based solely on image data was developed as a prototype for Army users, especially those users processing data over denied, restricted, or difficult access areas. The model was to be replicable, easy to use, and free of any user bias or subjectivity.

A future research goal will be to evaluate the sensitivity of this model against actual ground truth to see how well the model is performing and to evaluate imagery analyst contributions to final map output reliability. Use of ground truth data is not intended to become part of any future model extensions. Ground truth information is to be used solely to verify and validate our present model that uses only image source data in determining reliability.

Methodology

Project Site

The project site is a 3- by 4-kilometer area-of-interest inside the fence line of Fort A.P. Hill, located approximately two miles north of Bowling Green, Virginia. The site is considered upper coastal plain and is covered by a mix of upland and bottomland deciduous and coniferous forest, grassland, and urban built-up area. The installation is used extensively for U.S. Army Reserve training.

Source Material

Image source was acquired that was representative of data available to Army terrain analysts. SPOT XS multispectral imagery for June 1996 was collected. An orthorectified true color air photo mosaic compiled at 1:6000 scale for a January 1996 winter acquisition was also used. Ground truth data were acquired in June and July 1997. One hundred-seventy field sites were visited and detailed attribute information was recorded. Geographic information system (GIS) terrain data also were available for review.

Image Processing Software

Three COTS image processing packages were reviewed for their functionality in reliability mapping: IDRISI version 2.0, ERDAS Imagine version 8.3, and ENVI (The Environment for Visualizing Images version 2.6). The more established classification tools resided within all three packages, while newly developed uncertainty mapping capabilities are resident within IDRISI, but with *a priori* conditions necessary to maximize these functions. Three classic image processing techniques (Maximum Likelihood, Minimum Distance to Means, and Parallelepiped) offer classification probabilities if *a priori* knowledge is definable/available and incorporated into the models. A brief discussion of each of these tools follows, or may be reviewed in Avery and Berlin (1992):

Maximum Likelihood-- based on a probability density function associated with a particular "training site" signature. Training sites are image pixels that are pre-assigned into a terrain class by an image analyst. Pixels are assigned to the most likely terrain feature class based on a comparison of the *probability* that they belong to each of the remaining signatures being considered. The basic equation assumes that all classes have equal probability for pixels to be assigned to.

Variations on the equation, such as those that occur within ERDAS Imagine, allow for the analyst to override the equal class probability and to subjectively assign probabilities for each class that sums up to 1.0.

Minimum Distance to Means— based on the mean reflectance of each band for a signature, pixels are assigned to the class with the mean closest to the value of that pixel. To account for differences in the variability of signatures, minimum distance to means allows band-space distances to be normalized. It is commonly used when the number of pixels used to define signatures is very small or when training sites are not well defined. Within-class variance is not considered for this technique.

Parallelepiped— based on the minimum and maximum reflectances determined for a signature on each band. To be assigned to a particular class, a pixel must exhibit reflectances within this minimum – maximum range for every band considered. The parallelepiped procedure is potentially the least accurate of these three standard-bearers.

Signature separability can be measured statistically. The greater the distance between the means of each signature file, the greater the ability for equations such as Maximum Likelihood to correctly classify an image. This suggests that after analysis of the separability between individual terrain signatures, an analyst could apply this value into a probability coefficient.

Maximum likelihood is a supervised classification technique that examines the probability for each and every pixel signature within an image array to best be categorized within a most analogous, previously defined, feature group. The feature groups were the topographic features previously "trained" by the image analyst. The better the job completed by the image analyst at defining training sites, the greater the chance for an acceptable maximum likelihood derived output product. With poor or unreliable training sites one should consider the minimum distance to means technique. With variability in feature classes, maximum likelihood classifiers can interpret image pixels and classify them into correct feature classes.

Supervised classification means that some *a priori* knowledge has been "value-added" to the image to better allow the software to automatically characterize the image features. This knowledge may be acquired in many ways to include the use of maps, photos, site visits, discussions, other imagery sources, and text. Signatures are created from the imagery by training on areas that appear to be as homogeneous in cover type as possible. If one is confident in characterizing the feature found at a known location, then that location on the image may be "trained" as being that identified feature type. Once enough features have been geographically identified and located on the image space, supervised classifications techniques will look at the signatures of the trained pixels and search for analogous pixels within the image. The result is an image that has been better characterized by teaching the image "what-is-what" in the image space. The more ancillary data sources available for interpretive assistance to the analyst, the more likely the analyst is correctly training the image pixels for feature identification.

All properties on earth have measurable reflectance characteristics. In the case of living organisms, these signatures may change based on the time of year (Verbyla, 1995). Reflectance values may be acquired by remote sensing platforms and stored as digital numbers within an image array. In a perfect world scenario, each digital number (DN) or spectral value would represent the correct feature on the ground that has the corresponding reflectance value. Many impediments stand in the way of this one-to-one correlation:

> First, atmospheric conditions affect the ability to collect a pure, unaffected ground signature of features.
>
> Second, mixed pixels, or areas of non-homogeneous ground cover, present an averaging of reflectance characteristics. The net result is a spectral signature or digital number that is not necessarily representative of the real ground conditions.
>
> A third consideration is the season for which the data are collected. A spectral signature of a vegetation species will change dramatically over the growing season (Verbyla, 1995).

Given these spectral signature constraints, one may still successfully employ the power of ground feature signatures in characterizing a landscape. Ground truth collection for an area of interest is extremely valuable. Ancillary data are crucial (texts, maps, photos). Shape, texture, tone, orientation, pattern, and signature are all interpretive tools available to the trained image analyst. However, it is a spectral signature that offers the greatest potential for regional, automated interpretations of an image.

Layer Versus Pixel Reliability Mapping?

There are several ways in which reliability of terrain data can be considered. One way addresses terrain data *layers* with a reliability score assigned to each terrain theme. For example, the theme for vegetation may be divided into forest types pine, hardwood, and mixed, but the entire vegetation layer is scored with a single reliability regardless of forest type. A second way to focus on terrain data reliability would be to consider individual features within a terrain layer, and this may be done in a vector, object, or raster-based geographic environment. In the raster environment, which is most convenient for imagery-based terrain feature extraction, individual pixels may each contain reliability score. For example, a pine forest type within the vegetation layer could be the most accurately classified of the three forest types. Pine forest pixels might retain higher confidence, or reliability in the classification, than the hardwood or mixed forest. A method that combined terrain layer (or theme) *and* pixel (or within-theme) reliability was selected for investigation.

Signature Training-Set Development

SPOT XS imagery and a high-resolution photo mosaic imagery were imported. Geographically linking the two image products together was possible after the two products were projected to the same coordinate system (i.e., WGS 84). Side-by-side display of a SPOT scene and photo mosaic with geographic linking permits identical cursor orientation within each image space and facilitates the training signature development. This type of direct geo-linking of data sets can be foreseen for an Army analyst working with national assets and a commercial multi-spectral image source such as SPOT, Landsat, or IKONOS. Even without geo-linking, the process of signature development is not difficult when there are sufficient photo-identifiable cultural and/or natural features within the image space for the photo analyst to use for registration.

Photo interpretation of the Fort AP Hill photo mosaic resulted in the assignment of eight terrain class training signatures that were to be developed on the corresponding SPOT scene.

To minimize any chance that registration between image sources could negatively affect the training signature selection, only pixels that originated near the center of terrain theme polygons were selected; hard edges and ecotones were avoided. The eight terrain classes readily identifiable from the mosaic were

Pine Forest;
Hardwood Forest;
Mixed Forest;
Grass;
Urban/Built-Up;
Pond/Lake;
Stream/Drain; and
Road.

Despite continuing difficulty with delineating a stream/drain theme, all eight classes were selected for classification as they match up with specifications for terrain data as dictated by the National Imagery and Mapping Agency (NIMA) Tactical Terrain Data (TTD) and Feature Foundation Data (FFD) requirements. Each terrain class was defined by selecting five polygons with continuous pixel size totaling five or more each. The total number of "training" pixels per terrain class to be used later within a supervised image classification algorithm was approximately 100 to 150.

Signature Separability

Training pixel histograms offer a revealing evaluation tool for determining the spectral separability of imagery-derived terrain classes. A subjective approach is to plot all histograms atop one another and to visually evaluate the overlapping classes. Terrain classes that overlap will have greater difficulty distinguishing pixels that are appropriate for those classes. An example of a typical overlapping terrain class pair are pine and mixed forest, as the pine theme is obviously recognized as a component of the mixed forest signature. An analyst may look at a histogram of all the training pixel classes at once to get an understanding of the overlap to be expected between particular terrain classes.

ERDAS Imagine provides the user with a contingency table that reviews training pixel signature separability. Training signature separability was determined using the Mahalanobis distance decision rule, returning total number and percentage of training pixels classified as expected for each terrain class. Pixels with signatures that overlap, or are confused with similar terrain theme signatures, are misrepresented in terrain classes for which they are not intended. Percentages of training pixels classified as expected into the eight terrain themes were recorded and saved. This method is replicable and objective.

The computed contingency table percentages per terrain class are considered to be representative of a best-case scenario for classification since training sample pixels were specifically chosen by an image analyst because of their homogeneity and geo-linked match to a photo mosaic. This suggests that the entire multi-spectral image domain for the project site should be expected to only meet, and not exceed, the individual terrain theme reliability percentages unless the training signatures are adjusted. Accordingly, in development of a reliability methodology, training sample contingency table percentage values are considered as each terrain layer's maximum achievable reliability. Individual image pixels subsequently classified into a particular terrain layer would never achieve a reliability measure that exceeded an overall terrain layer reliability score.

Pixel Distance to Means Processing
Mahalanobis distance supervised classification was used to process SPOT pixels over the AP Hill study site. An important by-product available from a Mahalanobis distance method is a distance map. The distance map computed was a one-band, 32-bit continuous raster layer in which each data file value represents the result of a spectral distance equation, such as Mahalanobis. The equation for the Mahalanobis distance classifier is (ERDAS, 1999)

$$D = (X - M_c)^T (Cov_c^{-1}) (X - M_c)$$

where

D = Mahalanobis distance
c = a particular class
X = the measurement vector of the candidate pixel
M_c = the mean vector of the signature of class c
Cov_c = the covariance matrix of the pixels in the signature of class c
Cov_c^{-1} = inverse of Cov_c
T = transposition function.

The pixel is assigned to the class, c, for which D is the lowest value.

Unlike minimum distance and parallelepiped algorithms, covariances are computed and used for the Mahalanobis algorithm to standardize all the variables to the same variance. The Mahalanobis technique relies on parametric, or normally distributed data, within each input band of spectral data. Upon visual histogram examination, the spectral bands were deemed normally distributed.

Mahalanobis classification depicts all terrain themes in a composite graphic that permits examination of within-class pixel reliability through an ERDAS Imagine command: cursor-inquire-mode. However, the terrain themes may be analyzed more effectively if segmented from one another. Segmentation is accomplished using Imagine's <Image Interpreter/Utilities/Mask/Recode> functionality. Mahalanobis distance determines a statistical distribution of the pixels within a terrain class by computing a distance to class means unit of measure. An image analyst can select any pixel from the on-screen image domain and determine its statistical location (or distance) from the mean of its terrain class.

Histograms of terrain class distance values that have an exceptionally long tail away from the mean are an indication of pixel values with widely disparate reliability. Knowledge of a pixel's statistical location about the class mean is very useful information from which to assign a reliability score that assesses the confidence of each pixel's assigned classification category.

Results and Discussion

Terrain Class Pixel Thresholding
Distance images created from Mahalanobis have a Chi-square distribution, not a normal symmetrical distribution. Pixels with distances at the tail of the distribution represent pixels that are most likely misclassified and also appear to represent isolated pixels in the image space (Figure 1a). A cutoff point along the tail was both visually determined and computed statistically by using a Chi-square maximum distance value computed for a user-defined 95% confidence level. This level may be interactively adjusted depending on the desired confidence level. A visual and statistical approach to histogram tail removal showed that they approximate one another in final results. The final statistical approach selected to

minimize the outlier pixels along the histogram distribution tail was to use the Chi-square method, selected by choosing the "Threshold" command within Imagine's Spatial Modeler environment. This method is interactive and allows for changing of the confidence interval by the user. A combination of the "Clump" and "Sieve" commands was initially selected for statistical removal of individual, or small isolated contiguous pixels (outliers), but this approach proved ineffective due to lengthy processing time and inadequate user control over the process as compared to the "threshold" technique.

Pixels that remain after thresholding (removing the distribution tail), along with corresponding distance values (Figure 1b), constitute the range for new minimum and maximum distance measurements in a continuous floating point data structure. With a final distance measurement range defined, formulas may be written and applied against the individual pixel distances. An algorithm was written using the "Conditional" model developer that recalculated terrain class distance values into normalized pixel values with a new minimum of 0.001 and maximum of 1.0. This normalization of Mahalanobis distance values ensures a comparable metric for reliability scores across all terrain themes. Because the range of distance values in effect is always decreased by the threshold command, maximum Chi-square values represent outliers with highly suspect pixel classifications. Removal of pixels having the greatest distance values ensures that the normalization of the remaining pixels returns a reasonable approximation of the original distance values. The threshold command was critical, therefore, to the normalization process.

The algorithm developed for normalization of Mahalanobis distance values was further refined. Normalized distance values for each pixel were multiplied by their respective terrain class reliability score, computed earlier in the processing as the training sample contingency table percentage. Contingency table percent is easily converted to a value between 0.01 and 1.0, with value 0.01 signifying the maximum distance to class mean and 1.0 representing the exact class mean. Normalized pixel reliability values are multiplied with the overall terrain class reliability percentage, therefore all pixels assumed a floating point value between 1 and 100 percent. The following formula is an example of a computation for normalizing a pine forest pixel, where the pine terrain layer value was computed earlier from a training sample contingency matrix with score *0.8159*. This value changes for each terrain layer.

EITHER 0 IF <filename = 0> OR *0.8159* * {GLOBAL MAX <filename> -- <filename> / (GLOBAL MAX <filename> - GLOBAL MIN <filename>)} OTHERWISE

These new pixel values represent the terrain classification reliability. There is a potential to overstate the degree of confidence one could place on continuous data reliability scores at the pixel level. An analogy might be the erroneous practice of carrying significant digits out beyond that which is mathematically supportable. Are continuous reliability data scores really needed, or is a degraded qualitative format acceptable (e.g., poor, acceptable, good)? That question is probably best answered by the end user. Reliability information should probably not be degraded into categories because the original information is then essentially lost forever. However, the visual representation of the data could be more easily depicted by a reclassification without permanent adjustment to the data themselves. For example, simple cartographic presentation of the colors red, yellow, and green can be used to represent pixels considered of poor, acceptable, and good reliability. Development of a user interface to facilitate the re-classification and display of only those pixels of user-defined reliability is achievable within current image processing software packages.

Visual Representation of Reliability

Useful representation of reliability was examined using several approaches. The first attempt was to display a full continuum, or gradient, of certainty for a terrain theme; reclassification

into categories was not attempted at this point. A full spectrum of 256 colors is available within the computer palette and the result is a product that is very difficult to comprehend. The second attempt was a cartographic improvement to the first design, where the continuum of colors chosen to represent distance values was consolidated into three groups, as described in the previous red-yellow-green stoplight color approach. Grouping of pixels into the three categories was accomplished by a visual review of the raster attribute editor table for distance values and a manual thresholding of the pixels into recoded groups. This technique resulted in a map product that was easily produced and readily comprehensible to the user. The third and last approach selected for visual display was to use three-dimensional representation of reliability where distance from class means was assigned to the z-values and geographic location of the pixels was plotted in cartesian coordinate XY space. When plotted, pixels farthest from the mean value were shown as the tallest vertical spikes in the image space. "Flat terrain" represented pixels very near to the mean. This product was deemed to be an ineffective alternative in conveying terrain class reliability.

Future in Imagery-based Reliability Mapping

Techniques for image classification are changing. Emphasis on improved classification analysis can be seen in packages such as IDRISI where Bayes and Fuzzy analysis techniques are available. These newer tools are not nearly as mainstream as maximum likelihood or minimum distance but are emerging as viable complements (Foody, 1996). Geostatistics for image processing of land cover is also an emerging and promising solution. Each of these techniques considers data uncertainty as an important output. Transition of these certainty data to the software user into a useable geographic format is critical. Imagery-derived terrain data must be GIS supportable and a measure of their reliability is necessary. Probabilistic model output is possible if you start out knowing the confidence in the data and in the model itself. New models and methods for propagating terrain reliability will be needed in the future.

Conclusion and Summary: Topic 1

Tools resident within a COTS image processing package were flexible and functional enough to permit development of a terrain reliability model that did not demand ground truth. Formula development and pixel computations were completed within the spatial modeling environment. Formulas developed for this model are not believed to be specific to a geographic region or terrain data set. This will be determined in future model testing against ground truth.

The model developed for this project was not overly rigorous or abstract in nature. It was designed to be simple and easy to understand. As desired in the initial goal, the entire model process is replicable, easy to use, and free of any user bias or subjectivity. Terrain layers are still conventionally derived and may look identical to previously compiled terrain data, the only significant difference being the value added information detailing pixel reliability. This reliability information may be kept invisible to the terrain data user as simply pixel background information (raster attribute data) or it can be made readily apparent through creative cartographic display. Users who prefer to display the terrain reliability information have tools available within COTS image processing software to display the data at self-determined measures of reliability. Whether displayed or not, reliability information can be there when needed.

Terrain class pixel reliability may be integrated into decision analysis models. The confidence that decision-makers have in the decision analysis models will most clearly be affected by the reliability of the input terrain data. Terrain reliability integrated into decision models is compounded when more than one GIS layer of terrain data is considered.

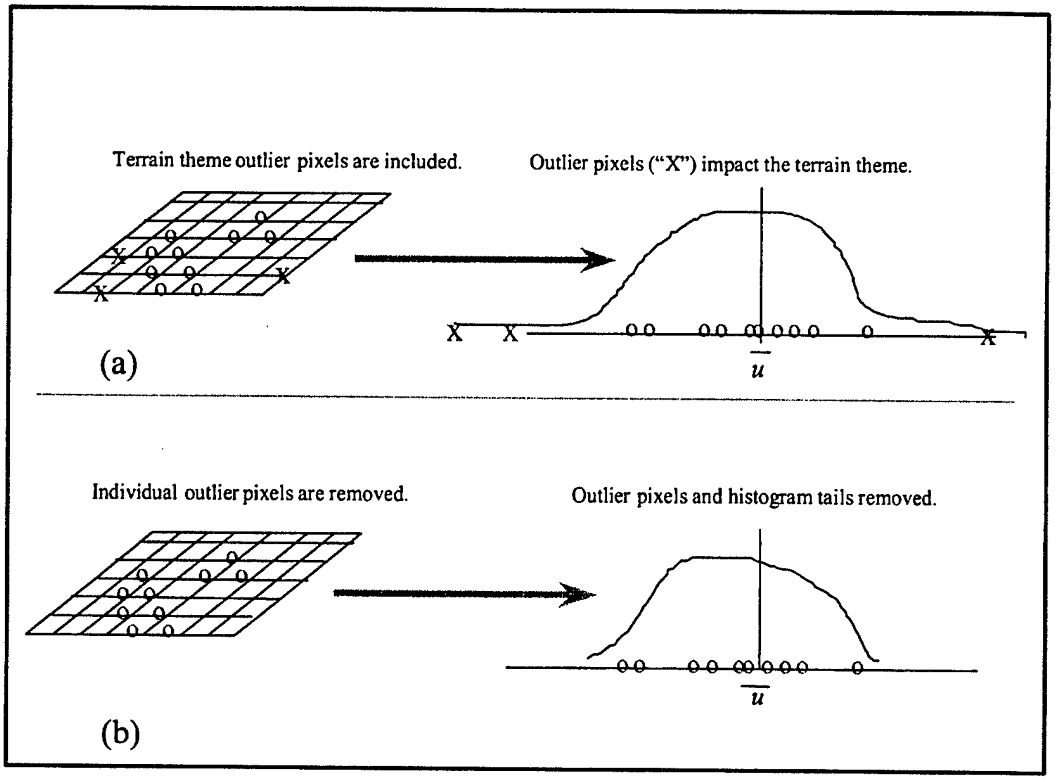


Figure 1. (a) Pixels classified within a terrain theme class to which they fall greater than three standard deviations from the class mean. (b) Pixels greater than three standard deviations from the mean (outliers) have been pruned from the class set, thereby decreasing the spectral range of the class, tightening the spectral signature, and increasing the chances for spectral separability from spectrally similar terrain classes.

Topic 2. Incorporation of Imagery-Derived Reliability Data into Tactical Decision Aids

Background

A tactical decision aid can be roughly defined as an initial and general guide to a commander in better understanding the battle conditions and environments and in making short-term combat decisions. It minimizes the difficulties of making combat decisions that commanders face every day. Over the years, many tactical decision aids (TDAs) have been developed and integrated into fielded systems such as the Digital Topographic Support System (DTSS) residing on the Combat Terrain and Information System (CTIS). Helicopter Landing Zone (HLZ) and Bivouac Sites/Assembly Area (BIV) TDAs are helpful in identifying suitable areas for landing helicopters and establishing camps, respectively. However, they are products that do not take into account reliability of source data. Users have requested knowledge of TDA product reliability and suggest a need for propagation of uncertainty through the spatial model decision-making process. Resultant output would be a map product that adequately portrays reliability to the user community.

Improved product quality depends not only on accuracy and precision but also on how products incorporate uncertainty. Numerous terrain themes can be used in TDAs: elevation, soil, vegetation, slope, drains, transportation, natural obstacles, etc. Every tactical decision aid requires some combination of themes of terrain data as inputs. For example, the HLZ TDA requires soil, slope, and vegetation as inputs. Supervised classification has been used to extract desired features from within an image source for use as data input to TDAs. However, as discussed in Topic 1, supervised classification of remotely sensed imagery will inevitably introduce data uncertainty in the terrain classes themselves and within the pixels that constitute the various classes. How does this uncertainty contribute, if at all, within the modeling environment? Currently, it is not a factor.

Objective

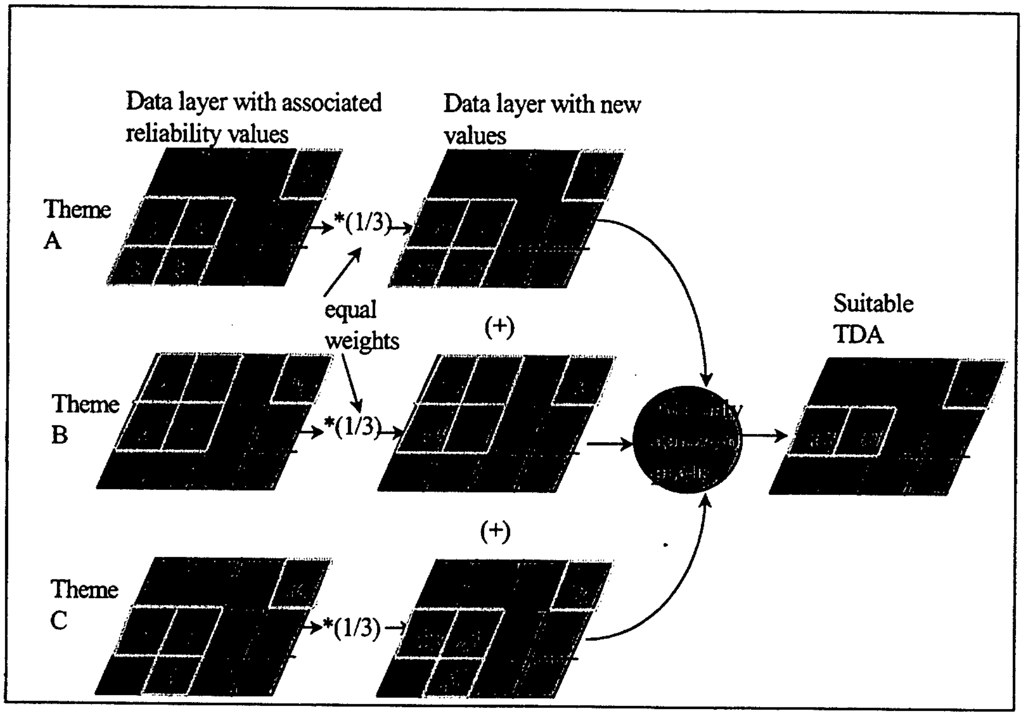
A method for reliability assessment and representation is needed that adopts simple-to-implement and easy-to-understand logic. Accordingly, the purpose of this study was to develop a nontechnical methodology that used individual pixel reliability and, subsequently, demonstrated the propagation of this reliability through a spatial tactical decision model. Pixel reliability computed from Topic 1 was regarded as the starting point for this effort.

Methodology

To integrate reliability pixel values computed earlier into a tactical decision aid, the difficult initial step is to determine the importance, or weighting, of each thematic terrain data layer in deriving an adequate product. A method of accomplishing this step is discussed in Topic 3: Terrain data requirements for HLZ and BIV. Users may have little knowledge as to which terrain layer is most important in contributing to a decision aid. Therefore, it would be reasonable to assign equal weighting to each terrain theme as a default value. For simplicity, an example TDA is described that has

three thematic data layers as inputs. Two approaches to mapping terrain layer reliability were evaluated: linear combination and fuzzy classification. Linear combination is computed by multiplying the reliability values of every pixel within the respective thematic data layers by 1/3, then spatially summing all non-zero reliability values across the three new value layers (Figure 2). A pixel is assigned to class *m* along with a value measuring its degree of reliability to belong to class *m,* as computed by Mahalanobis distance. Mahalanobis distance is not a probability score, nor is it a measure of chance for terrain class *m* to be found at a particular pixel. This approach we have taken is very similar to the work of Zhu (1997), where measures of uncertainty are provided with class assignment.

Figure 2. Linear combination for overall reliability values for TDAs.



In Figure 3, fuzzy classification method is illustrated with the assignment of a pixel to more than one class. Generally, fuzzy classification is a methodology to assign a pixel to each of a set of classes (more than one class) and to indicate the degree to which the pixel belongs. Fuzzy logic models the degree to which a pixel belongs to a class, otherwise known as the degree of membership. More specifically, pixel reliability assignment comes from the lowest value of the three thematic data layers applied to TDA.

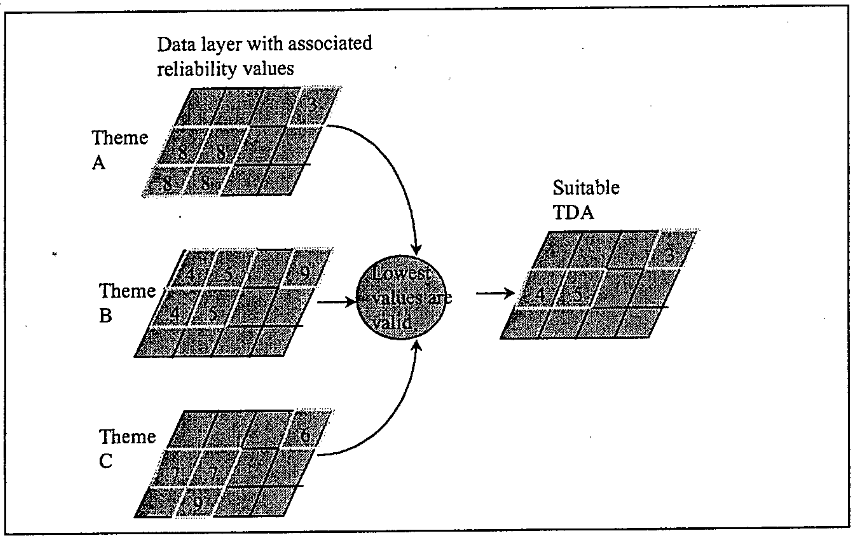


Figure 3. Fuzzy logic for overall reliability values for TDAs.

If thematic data layers are determined by human terrain analysts to play unequally important roles as inputs, a user interface should allow the user to set the weighting that is considered as appropriate to each of the themes and relative to others. Topic 3 provides information on how to more objectively define terrain theme weights. Knowledge of terrain theme contribution to model output for various physiographic domains may better enable a user to weight individual terrain themes over others. Altering the weighting of terrain themes results in different output products for comparison and analysis.

We apply the same linear combination method by multiplying the reliability values of every thematic data layer with the user-defined weights, and then spatially summing the new non-zero reliability values. The weighting range should be 0.0 to 1.0, so that addition of individual terrain theme reliability values together sums up to a maximum overall reliability value of 1.0. While this type of reliability is not probablistic in nature, it does provide a readily computable metric. Once computed, the complete range of reliability metrics can be divided into three categories, for example, representing good-, fair-, and poor-reliability, or any other user-specified number of categories.

Results and Discussion

Implementation of Reliability into a Sample Helicopter Landing Zones Model

A simplified HLZ TDA requires thematic data layers for soil, slope, and vegetation as inputs. Suitable conditions required for landing are as follows:

- Soil is gravel or sand
- Vegetation is barren, pasture, grassland, or dry agriculture
- Slope is within 0 to 3%.

Pixels from soil, slope, and vegetation data layers that meet the above selection criteria are combined together by using the Boolean AND (INTERSECT) operator within an Arc/Info GIS environment. Pixels that do not meet the selection criteria are ignored and classified as unsuitable for the particular HLZ TDA. Suitable pixels each carry along an associated reliability value computed earlier. It is implied at this point that the imagery-derived method for assigning pixel reliability has been used prior to this step to define soil and vegetation pixel reliability. Slope reliability values developed for this experiment were not computed from the method described in Topic 1 as slope was computed from integer data, and vegetation and soil reliability was computed from categorical (class) data. Reliability values for all pixels from each terrain theme are multiplied by a user-defined weight for that theme to obtain newly derived pixel reliability values. The non-zero pixel values are summed across all terrain themes and a "Suitable HLZ" product is generated with commensurate reliability values (Figure 4). It should be clear that terrain themes that are weighted highest (e.g., vegetation at 0.5) contribute more to the total reliability score than lesser-weighted themes.

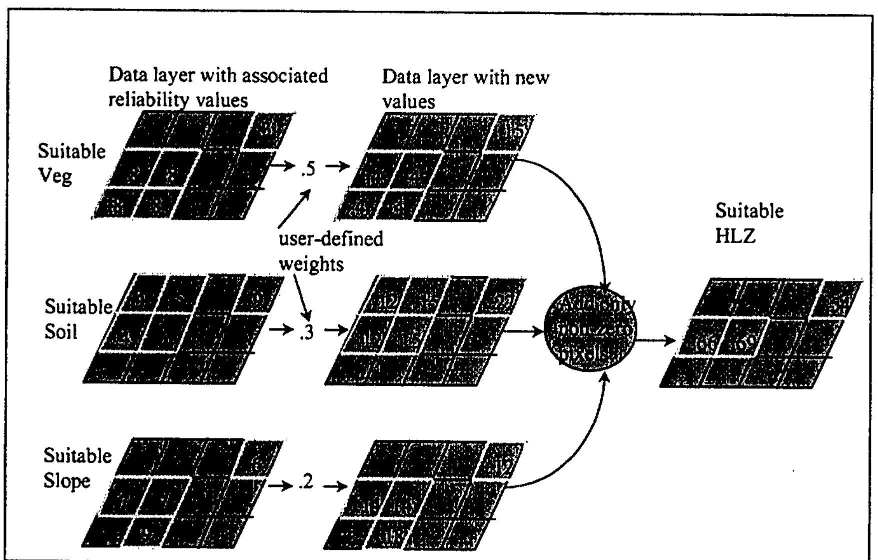


Figure 4. Linear combination of overall reliability values for HLZ.

As depicted in Figure 4, the vegetation, soil, and slope layers are in raster format and have 5, 5, and 4 suitable reliability pixels, respectively, associated with each of the original three terrain themes. These values were multiplied by weights assigned by a user to each terrain theme, in this case, 0.5, 0.3, and 0.2, to obtain the individual reliability value after taking the weighting into account. To obtain the overall reliability for suitable HLZ, we would spatially sum the new reliability values of the vegetation, soil, and slope layers and note that only pixels with non-zero values would be summed. Accordingly, the result was an HLZ product that depicts three suitable pixels representing a known location and size. Pixels with large reliability values are most desirable. At this point, we could qualitatively divide the overall reliability range into a user-specified number of categories such as good, fair, and poor. Thus linear combination method with weighting is the technique recommended for incorporating reliability data into tactical decision aids.

Code Suggested To Be Used in Arc Macro Language

Example Arc/Info Arc Macro Language (AML) code is provided that illustrates the exact coding mechanics involved in implementing reliability into a TDA. Although the code provided is explicitly for an HLZ model, the design should be extendable to any suitability-type model.

- Pre-reliability code executed in an Arc module for definition of suitable HLZ binary product:

```
...
intersect %soil_for_hlz_cov% %veg_for_hlz_cov% %soil_veg_cov%
intersect %soil_veg_cov% %sel_slp_0_3_cov% %hlz%
...
```

- Post-reliability code executed within the Arc/Info Grid module for definition of suitable HLZ product displaying a range of suitable values:

```
...
&sv weight_for_veg 0.5                              /* weights from figure 3
&sv weight_for_soil 0.3
&sv weight_for_slp 0.2
grid
%weighted_veg% = %veg_for_hlz_cov% * %weight_for_veg%
%weighted_soil% = %soil_for_hlz_cov% * %weight_for_soil%
%weighted_slp% = %sel_slp_0_3_cov% * %weight_for_slp%
%soil_veg_cov% = %weighted_veg% + %weighted_soil%    /* a zero value here is equivalent to
                                                        a NODATA value in Arc/Info
%hlz% = %soil_veg_cov% + %weighted_slp%
...
```

<u>Conclusions and Summary: Topic 2</u>

A linear combination method, with weighting either subjectively user-defined or obtained from a method such as described in topic 3, is a simple technique recommended for incorporating reliability data into tactical decision aids. The method is replicable, portable, and easily implemented. Data reliability propogates through the model and the resultant map output is a satisfactory representation of reliability from which a user may make a more informed decision. Fuzzy mapping returns a worst possible case interpretation of the data reliability that can be misleading. One might get the decided impression that the data are not worth using. Linear combination method seems to be a bit more "even-handed" in its results. Given a level of disparity in output-pixel reliability, decision-makers are equipped to make better informed responses to tactical decisions. Model output reliability as directly affected by data input reliability provides critical information that may determine the anticipated success or failure of a mission.

Topic 3. Terrain data requirements for Helicopter Landing Zone and Bivouac Models

Background

Army terrain model sensitivity to missing data is not fully understood prior to generation of a final map output product. Existing tactical decision models accommodate missing data, but at a relatively unknown cost to the overall reliability of the final map product. Site suitability models, in particular, such as HLZ and BIV sites, might deliver a reasonable output product from a subset of the requested terrain input data. Determining an appropriate subset of input terrain data with which a user could still expect a reliable model end-product is addressed in the following pages.

Demand for digital data products continues with the National Imagery and Mapping Agency (NIMA), as developer of standardized topographic data, simply unable to keep up with the global demand. Soil data are exceptionally rare, and vegetation is temporal in nature. Terrain models have been designed to use all pertinent data in generating analysis products, and models that compute a map product with partial input data sets should provide a measure of the resultant map confidence. This has historically not been the case.

Objective

The objective of this research was to understand the level of contribution of terrain data layers, both individually and in combination. The research also was designed to look for possible climate region correlation of terrain data layers at different study sites.

Methodology

An experiment was defined in which terrain data for two tactical models were evaluated and measured for their overall contribution to the final map products. Measuring the contribution to the final map products enabled the prioritization of data requirements for each of the two models, which is important due to finite resources (e.g., time, money, data, personnel). When evaluating data needs for processing a specific terrain analysis model, the most crucial data should be the first data considered for collection and processing.

Study Sites

Three geographically unique study sites were selected for evaluation:

1. Korea (near Changchon-ni)
2. Ft. Hood, Texas
3. Camp Pendleton, California.

Respective climate regions for these three locations, according to the Koppen-Geiger Climate map, are Cwa, Cfa, and Csb[1] (Tromblay, 1953). Additional geographic locations within these identical climate regions were then selected as test sites.

Climate Zone Portability

To assess the potential portability of our research findings, these testing site data were to be compared to the original study sites at Korea, Fort Hood, and Camp Pendleton to determine if there was portability of data requirements from one climate region to another identical region. It was hypothesized that if the data priorities were reported as equivalent for each climate region "pair," general rules could be established for priority of data acquisition. The examination of climate regions, as opposed to the individual topographic variables, sought to look for a general solution to the problem.

The test sites for Cwa and Dfa climate regions were both located in Korea (site names Kor_34202 and Kor_35151, respectively). No test site containing standard digital feature data was available within an alternative Csb site to the Pendleton area. Accordingly, evaluation and test-site pairs were as follows, with a Pendleton pair excluded:

Study site	Test site	Climate Region
1. Camp Pendleton	*None available*	Csb
2. Ft. Hood	Korea (Kor_35151)	Cfa
3. Korea	Korea (Kor_34202)	Cwa

Data Source

The data selected for investigation came from Interim Terrain Data (ITD), produced as a standard terrain feature product by NIMA, and Digital Topographic Data (DTOP), a prototype NIMA feature data base.[2] Imagery data also was analyzed over Camp Pendleton for an alternate follow-on investigation. The following table lists the study and test sites and the data used for their analysis:

---

[1] The Koppen-Geiger Climate map defines Cwa, Cfa, and Csb as follows:

First letter	C: sufficient heat and precipitation for growth of high-trunked trees.
Second letter	f: sufficient precipitation in all months.
	s: dry season in summer of the respective hemisphere.
	w: dry season in winter of the respective hemisphere.
Third letter	a: warmest month mean over 71.6°F (22°C).
	b: warmest month mean under 71.6°F (22°C). At least four months have means over 50°F (10°C) (Tromblay, 1953).

[2] Interim Terrain Data (ITD) and Digital Topographic Data (DTOP), a prototype data base, provided the feature data used for the HLZ and BIV models. ITD has six thematic layers: Slope/Surface Configuration, Soil/Surface Materials, Vegetation, Surface Drainage, Transportation, and Obstacles, each of which is divided into several features and attributes. DTOP has essentially equivalent terrain feature and attributes.

Sites	Terrain Feature Data	Imagery	Area
1. Camp Pendleton	DTOP	Landsat TM	451 km^2
2. Ft. Hood	ITD		677 km^2
3. Korea	DTOP		744 km^2
4. Korea (Kor_35151)	ITD		661 km^2
5. Korea (Kor_34202)	ITD		741 km^2

Locating several test sites within geographic areas climatically designated Cwa and Cfa was not possible due to the limited ITD coverage. Each of the five sites evaluated was approximately equivalent in total size to one 1:50,000-scale map sheet.

Two terrain analysis models were selected for experimentation:

1. Helicopter Landing Zones;
2. Bivouac Areas.

Both of these site-suitability models were rewritten using Arc Macro Language to accept incomplete data sets of ITD and DTOP alike, and they subsequently ran successfully within an Arc/Info geographic information system (GIS) environment. The Arc Grid module was selected for analysis. Proximity-to-transportation is also considered in some models, but because it was not implemented within the TEC-fielded DTSS system, it was not incorporated into software re-engineering for this project.

The following features and attributes were used as selection criteria for Helicopter Landing Zones:

Suitable:

1. Soil – all gravel and sand;
2. Vegetation – barren, pasture, grassland, dry agriculture;
3. Slope – 0 to 3%.

Suitable with caution:

1. Soil – all gravel and sand;
2. Vegetation – barren, pasture, grassland, dry agriculture;
3. Slope – >3 to 10%.

An Arc/Info version of the HLZ algorithm was re-coded with the Defense Mapping Agency Feature File (DMAFF) and Feature Attribute Coding Catalog (FACC) attribute schemes. Helicopter Landing Zone map output was designed to show the results from the various combinations of soil, slope, and vegetation.

The following features and attributes were used as selection criteria for BIV:

1. Soil – a Smooth Surface,
   or a Smooth Bare Rock,
   or Sand Dunes, Loess, Karst, Leteritic, Permafrost,
   or Stony Soil with Scattered Surface Rock,
   or Scattered Surface Rock,

or Alluvial Fans,
and with Dry Soil Moisture Content Conditions.

2. Vegetation – Coniferous, Deciduous or Mixed Forest during the summer,
and with Canopy Closure greater than or equal to 50%,
and Tree Height greater than or equal to 5 meters,
and Tree Spacing greater than or equal to 2.5 meters.

3. Slopes – 0 to 10%.

Bivouac map output was designed to show the results from the various combinations of soil, slope, and vegetation.

Statistical evaluation plan

A reference map was derived from a combination of soil, slope, and vegetation terrain data layers. Ground-truthing was not conducted on the model output. The output from all three data layers was considered to be the "reference map." Map output from these three terrain layers was assumed to be 100 percent accurate, or a benchmark upon which all other combinations of data were to be statistically compared. The Arc/Info Boolean operator AND (INTERSECT) was used on the three terrain data layers to create six data sets. They are Veg, Soil, Slope, Soil+Slope, Soil+Veg, and Veg+Slope. The analysis technique applied was to determine Kappa coefficients for each of these six data sets.

Kappa is a spatial statistics method that measures the randomness in the errors between the tested map output and the reference map output. The tested output layers and combinations were a) vegetation, b) soil, c) slope, d) soil + slope, e) soil + veg, and f) veg + slope. The following values constitute Kappa randomness values:

0.00 to 0.40 High randomness and therefore little correlation;
0.41 to 0.75 Moderate randomness and moderate correlation;
0.76 to 1.00 Little randomness and high correlation (Fleiss, 1981).

Kappa coefficients were calculated and assigned to each of six combinations of terrain data to see how closely the respective output matched the output from the reference map. The Kappa coefficients were prioritized based on their value. The highest Kappa values represented layers, or layer combinations, with the greatest contributions to matching the reference map. Therefore, prioritizing data layers was based on the highest Kappa coefficients.

In order to efficiently compute Kappa scores of comparison between multiple data sets, it was necessary to convert the terrain data layers from vector data to raster data. Vector data were converted into 30-meter raster data and then the raster data were analyzed on pixel-by-pixel basis, always comparing against the reference map.

Results and Discussion

Table 1 shows the Kappa coefficients for the HLZ model, which were obtained by conducting a quantitative analysis of each of the six subsets in comparison with the

reference data set. As would be interpreted by Campbell's work (Campbell, 1987), a Kappa value of 0.7722 such as was calculated for the veg + slope category for Korea, means that the accuracy of the classification is 77.22 percent better than that from random assignment of pixels to categories. Therefore, the larger the Kappa value, the larger the contribution of that data layer. Conversely, using the Korea study site again, a Kappa value of 0.0170 such as was calculated for soil means this terrain layer contributed little to explaining the HLZ output. Independently, and for all sites, each terrain layer exhibits a Kappa score of high randomness and little correlation.

Table 1. Kappa Coefficients Matrix for the HLZ Model.

Data	Korea (Cwa)	Kor_34202 (Cwa)	Ft. Hood (Cfa)	Kor_35151 (Cfa)	Pendleton (Csb)
Veg	0.1906	0.2055	0.1220	0.2634	0.2742
Soil	0.0170	0.0110	0.2468	0.0576	0.2000
Slope	0.1164	0.3823	0.0945	0.1630	0.4588
Soil+Slope	0.1836	0.4947	**0.6178**	0.3982	0.6517
Soil+Veg	0.2022	0.2857	0.5772	0.4005	0.3607
Veg+Slope	**0.7722**	**0.5985**	0.1587	**0.4422**	**0.6687**

Overall, Table 1 reveals a moderate-to-high dependence from the terrain variable combination of Veg + Slope for each Korea study site and for Camp Pendleton. Fort Hood relies on soil and slope as the most important variable combination as indicated by the 0.6178 value assigned. Individually, the terrain factors of primary importance were not consistent.

The prioritized order of Kappa coefficients of the individual data layers for vegetation, soil, and slope from test site Kor_34202 is *different* from the Korea (Cwa) study site, despite the fact that both sites are in the identical climate region. For Kor_34202, slope is the most important individual variable, whereas for Korea vegetation is the crucial single variable. Similarly, Ft. Hood and test site Kor_35151 have different terrain data priority and optimum data combination despite their identical climate zone (Cfa). Accordingly, climate zone does not appear to be a determining factor in data priority for HLZ modeling.

Soil was the most important individual variable (Kappa 0.2468) at Fort Hood. Slope was relatively insignificant as the area is uniformly flat (Kappa 0.0945). Vegetation was not particularly important as the majority of Fort Hood is devoid of dense vegetation (0.1220). There are only insignificant pockets of vegetation in the riparian zones and in the lower southwest corner of the installation. Accordingly, when vegetation was removed as an available variable, the percentage of HLZ area correctly identified as acceptable was not detrimentally impacted as it was for Korea (39.57% versus 3.56%).

Camp Pendleton has increased Kappa values for each of the individual data themes. Kappa coefficients are Vegetation Kappa 0.2742, Soil Kappa 0.2000, and Slope Kappa 0.4588. Slope alone shows moderate correlation to the reference map. The percentage of HLZ area identified by slope as suitable, with caution, rises to 75.42 percent. There is considerable landform variability at Camp Pendleton, easily visible

to a terrain analyst, helping to explain the importance of slope in locating suitable sites for landing helicopters.

Korea 35151 was a case study that illustrated there is not much difference in the correlation between the reference map and either soil+slope, soil+veg, or veg+slope. For this area of Korea, any combination of two of the three terrain variables would have yielded approximately the same results. Kor_34202 reveals a moderate Kappa score for veg+slope (0.5985) and a score of 0.3823 for slope alone.

A small additional experiment was run for the Camp Pendleton site. DTOP vegetation data were exchanged for vegetation data derived from Landsat Thematic Mapper, using an unsupervised Normalized Difference Vegetation Index. The result was an overall increase of 13 percent in the number of suitable HLZ pixels when using the imagery-derived vegetation data layer as opposed to using the more outdated DTOP vegetation data. There was a change in vegetation coverage area between the DTOP and imagery derived data layers with the newer imagery showing that there were fewer treed areas than the DTOP data showed years earlier.

Table 2 shows the Kappa coefficients for the BIV model, which were obtained by conducting a quantitative analysis of each of the six subsets in comparison with the reference map. Bivouac Kappa values for Pendleton are not shown because the condition for tree stem spacing greater than or equal to 2.5 meters, as required by the BIV model criteria, does not exist in the available DTOP for Camp Pendleton. Veg + Slope is once again the preferred combination of two terrain factors to be used for generating the most reliable BIV model in the Korea study area. A Kappa value of 0.8528 suggests that the accuracy of the classification is 85.28 percent better than that from random assignment of pixels to categories. The Fort Hood study area does *not* identify the terrain layer combination of Soil + Slope as the preferred data pair, as chosen previously for HLZ, but instead selects Soil + Veg as evidenced by the reported 0.5267 Kappa value.

Table 2. Kappa Coefficients Matrix for the BIV Model.

	Korea (Cwa)	Kor_34202 (Cwa)	Ft. Hood (Cfa)	Kor_35151 (Cfa)	Pendleton (Csb)
Veg	0.0605	0.1070	0.2605	0.1511	No data
Soil	0.0926	0.0186	0.3272	0.1444	No data
Slope	0.4236	0.3876	0.0689	0.3091	No data
Soil+Slope	0.8136	0.4891	0.4755	0.7370	No data
Soil+Veg	0.1066	0.1070	**0.5267**	0.1721	No data
Veg+Slope	**0.8528**	**0.9709**	0.3244	**0.8591**	No data

In Table 2, the prioritized order of the Kappa coefficients of the individual data layers for vegetation, soil, and slope of the test site Kor_34202 is similar to the Korea study site. Both sites are in climate region Cwa and suggest a Veg + Slope terrain data combination is most successful at representing accurate bivouac areas. A major misrepresentation could occur, however, if one was to use the successful Kappa results for Korea using Soil + Slope (0.8136) and apply these results as anticipated reliability for a terrain data set of Soil + Slope for Kor_34202, based on the fact that

they have an identical climate zone. Kor_34202 has a much lower Kappa coefficient for Soil _ Slope (0.4891) than would have been incorrectly predicted. Likewise, the study area pair Fort Hood and Kor_35151, located in climate region Cfa, do not agree as to recommended terrain data sets. They have completely different priorities both for individual terrain themes and combined terrain data. Fort Hood results are best for Soil + Veg (0.5267), while Kor_35151 results are best for Veg + Slope (0.8591). This finding is additional evidence that climate regions are not an effective controlling factor in determining critical terrain data priorities for individual terrain models.

The prioritized order of combinations of data depends mostly on the prioritized order of each individual data layer. The Korea study area clearly illustrates the importance of slope as the primary contributor to the model output. Slope and vegetation, in combination, return a Kappa coefficient of 0.8528, exhibiting high correlation and little randomness in comparison with the reference map. Fort Hood illustrates the lack of importance of slope in this area, much like for the HLZ model. Soil takes precedence among the individual data layers. Kor_34202 and Kor_35151 show the vegetation and slope combination as representing a 0.9709 Kappa score. In this geographic area, the addition of soil data adds little to improving the quality of the final bivouac suitability output. Lastly, Korea_35151 shows the vegetation and slope combination as representing 0.8591 and 0.8591 Kappa scores, respectively. In these geographic areas of Korea, also, the addition of soil data is of least importance in improving the quality of the final bivouac suitability assessment.

Conclusions and Summary: Topic 3

Arc/Info GIS algorithms were recoded for the HLZ and BIV models. Both models successfully ran without all requested terrain data layers.

General comments can be made regarding the study of terrain variables on HLZ output. First, in areas where there is significant vegetation, vegetation information is a critical terrain layer. Vegetation is a variable that remote sensing technology can exploit for rapid data generation. Not surprisingly, in areas devoid of significant vegetation (semi-arid land similar to Fort Hood), vegetation inclusion had little impact on HLZ output. These common-sense results can be applied to other areas of the world for terrain data prioritization.

Based on data analysis of the three Korean sites, data priority for both the HLZ and BIV models indicated Veg + Slope as the most critical terrain combination. With these two layers, there was moderate correlation to the reference map for the HLZ products (0.44 to 0.77), and high correlation with the BIV products (0.85 to 0.97). Soil data contributed very little to the final output reliability in this geographic area of the world.

Data analysis of HLZs for Camp Pendleton indicated that Veg + Slope were the most critical terrain parameters for portrayal of a moderately reliable HLZ product (0.67). The use of Soil + Slope presented a reasonable alternative as the Kappa value was only slightly less than that of Veg + Slope (0.65). Lack of needed terrain data

required to run the BIV model prevented the sensitivity testing of terrain data critical to Camp Pendleton.

Fort Hood placed greatest emphasis on soil data in determining both HLZ and BIV products. The data priority was to first select Soil + Slope (HLZ = 0.62, BIV = 0.53) or alternatively Soil + Veg (HLZ = 0.58, BIV = 0.48) as a close second choice. The combination of Veg + Slope did not produce a quality HLZ or BIV product as illustrated by low Kappa values associated with each terrain model. A synopsis of the research results for crucial individual terrain data needs is found in Table 3.

Table 3. Data priority for HLZ and Bivouac TDAs, identified by project study site.

Project Site	HLZ Data Priority	BIV Data Priority
Pendleton	1. Slope	N/A
(Csb)	2. Vegetation	N/A
	3. Soil	N/A
Ft. Hood	1. Soil	1. Soil
(Cfa)	2. Slope	2. Vegetation
	3. Vegetation	3. Slope
Korea	1. Vegetation	1. Slope
(Cwa)	2. Slope	2. Soil
	3. Soil	3. Vegetation
Kor_35151	1. Vegetation	1. Slope
(Cfa)	2. Slope	2. Vegetation
	3. Soil	3. Soil
Kor_34202	1. Slope	1. Slope
(Cwa)	2. Vegetation	2. Vegetation
	3. Soil	3. Soil

While perhaps intuitive in nature, general rules-of-thumb were gleaned from the results of this work:

1. The smaller the total area of pixels within a terrain data theme that are deemed suitable for BIV or HLZ, the more critical and important that data layer becomes. In other words, as a terrain theme became more restrictive, its contribution and importance grew.
2. The flatter, more gently sloping study area (Fort Hood), showed little importance on inclusion of slope data for either HLZ or BIV model output. Soil was most critical.
3. For sites with more diverse terrain (all except Fort Hood), soil became the least important terrain variable for HLZ and BIV. A combination of slope and vegetation was the most important terrain combination for HLZ and BIV.

An original goal of this research topic was to determine if there was a relationship between the worldwide climate regions and the terrain data priorities required by the HLZ and BIV models.

Climate zone did not appear to be a determining factor in data priority for either HLZ or BIV modeling. This is most likely because the climate regions are too vast in geographic area and the physiographic differences are great. Results from this work disagree with an investigation completed by Green *et al.* (date unstated), whereby they identify terrain data requirements for a model, also, but contend that climatic zones provided a reasonable geographic framework for the portability of model results.

This research provides evidence that there is a basis for recommending terrain data layer priorities crucial for adequate production of HLZ and BIV alike. Extrapolation of these research findings to other geographic areas of the world would have to be done with caution. Some type of "pairing" to analogous physiography (soil, landform, vegetation, slope) rather than climate may be a more plausible method for extending the above results to a new region of the world. Applying rules-of-thumb appears to be a sensible solution to data prioritization during instances when terrain information is not available and resources needed for their generation are limited.

Overall Conclusions and Summary: Topics 1, 2, and 3

Imagery-based terrain data are crucial for input into tactical decision models in the absence of available standard data. This is especially true for many vegetation and soil data feature types utilized in numerous Army-fielded models (see Appendix 1). Using COTS tools, a terrain layer reliability metric was computed using a training sample contingency matrix, and a within-terrain-layer pixel reliability metric was computed from a Mahalanobis distance classification tool. A spatial model was then prototyped to combine these two metrics into a single pixel reliability value that could be passed along with the terrain data as pixel-specific reliability scores. Pixel reliability was propogated through an HLZ model to test the generation of a sample map reliability product. Output can be categorized into high, moderate, and low-reliability zones to assist the decision maker in selecting most appropriate sites. Climate zones were not good predictors of terrain data requirements supporting HLZ or BIV models. Rather, physiographic regions may be reasonably considered as predictors for prioritization of HLZ and BIV input data requirements. Lastly, rules-of-thumb based upon similar physiographic study areas appear plausible and could be used to suggest terrain variables of highest required priority in fulfilling accurate HLZ and BIV modeling output. For example, in the relatively flat semi-arid study site of Fort Hood, soil data were the most important contributing variable, whereas vegetation and slope combined were the most important variables for HLZ and BIV models from the Korea study sites. Collectively, Topics 1 to 3 described a method for deriving reliability, integrating the resultant scores into a decision model, determining objective terrain theme weights to help drive sample decision models, and development of products that have immediate applied use by decision makers.

Appendix 1. Vegetation and soil data requirements that pertain to Army Tactical Decision Models (TDAs)

		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
VEGETATION																			
Marsh/Swamp																			
Marsh/Swamp-Type		X																	
Barren Ground		X	X	X	X														
Cropland			X	X	X														
Cropland-Type	Terraced, Shifting Cult, unk	X																	
Grassland			X	X	X														
Grassland-Type	Pasture, grassland w/trees	X																	
Scrub/Brush																			
Scrub/Brush-Density	Open-med Spacing, Med-dense	X																	
Trees-Type	Mix, Pine, Hard, Clearing	X		X		X	X	X	X	X	X	X	X	X	X				
Trees-Brush Understory	Sparse and Dense	X						X					X	X	X				
Trees-Canopy	25% categories						X	X	X	X	X	X							
Trees-Predominant	5m categorical	X						X			X	X	X				X	X	X
Trees-Stem Diameter	integer >0 to 900cm	X				X	X				X		X						
Trees-Stem Spacing	algorithm for min distance to nearest stem	X				X	X				X		X	X					
Trees-Foliage Height							X												
Trees- Penetrable/Impenetrable							X												
Vegetation Roughness						X													
SOILS																			
Soil Depth	< or > than 0.5m					X					X		X						
Surface Roughness	long detailed					X							X						
Soil Type	8 texture types/mixtures		X	X	X	X	X				X		X			X			
Soil Wetness Condition	dry, moist, wet						X				X		X						
Exposed Bedrock	long detailed										X								

Spreadsheet Legend of Model / TDA Names:

A. Vegetation
B. Drop Zone
C. Helicopter Landing Zones
D. Avenues of Approach
E. DMA Mobility
F. Cross Country Mobility (CCM)
G. Cover
H. Concealment
I. Observation and Field of Fire
J. Construction Resources
K. Key Terrain
L. NATO Reference Mobility Model II (NRMMII)
M. Bivouac
N. Aerial Concealment
O. Soil Trafficability
P. Point to Point Line of Sight
Q. Masked Area
R. Aerial Detection

References

Aronoff, S., 1982. Classification accuracy: A user approach. *Photogrammetric Engineering and Remote Sensing*, 48 (8): 1299 – 1307.

Aronoff, S., 1985. The minimum accuracy value as an index of classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 51 (1): 99 – 111.

Avery, T.E. and Berlin, G.L., 1992. Fundamentals of Remote Sensing and Airphoto Interpretation. Macmillan Publishing Company, New York.

Campbell, J.B., 1987. Introduction to Remote Sensing. Chapter 12, The Guilford Press, New York.

ERDAS, Inc., 1999. ERDAS Imagine image processing software, Image Analysis module.

Fleiss, J., 1981. Statistical Methods for Rates and Proportions. John Wiley and Sons, New York.

Foody, G.M., 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely-sensed data, *International Journal of Remote Sensing*, 17 (7): 1317 – 1340.

Green, J.G., Willoughby, W.E., Mason, G.L., Ahlvin, R.B., and Bullock, C.D., (date missing). Mission Specific Data Set Requirements for Wheeled and Tracked Vehicle Mobility Modeling.

Jensen, J.R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective. 2nd ed., Prentice Hall, Inc., New Jersey.

Slocum, K.R., Doung, V., and Moeller, J., 1997. Reliability mapping of imagery-derived terrain feature classes, USACE Surveying Mapping and Remote Sensing Conference, St. Louis, MO.

Storey, M. and Congalton, R.G., 1986. Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*, 52 (3): 397 – 399.

Strahler, A.H., 1980. The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment*, 10: 135 – 163.

Tromblay, J.P., 1953. Koppen-Geiger System of Climate Classification, after R. Geiger and W. Pohl. John Wiley and Sons, New York.

Verbyla, D.L., 1995. Satellite Remote Sensing of Natural Resources. Lewis Publishers, New York.

Zhu, A., 1997. Measuring uncertainty in class assignment for natural resource maps under fuzzy logic, *Photogrammetric Engineering and Remote Sensing*, 63 (10): 1195 – 1202.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
12-2002	Final	

4. TITLE AND SUBTITLE
A Suggested Methodology for the Computation of Imagery-Based Terrain Data Reliability and Implementation Into Tactical Decision Models

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Kevin R. Slocum, Ph.D

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Engineer Research and Development Center (ERDC)
Construction Engineering Research Laboratory (CERL)
PO Box 9005
Champaign, IL 61826-9005

8. PERFORMING ORGANIZATION REPORT NUMBER
ERDC/TEC TR-02-1

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Copies are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

14. ABSTRACT

The issue of data uncertainty and its impact on model output is discussed in this report under three topic areas. Modifiable methodologies for implementation of data reliability are provided for each topic. The first topic addresses the fact that new digital terrain data are created daily that directly supports tactical decision-making. It is common for decisions to be made without the full understanding of the contributing terrain data quality. Data are treated as spatially invariable in quality and devoid of any metric measuring the underlying certainty of feature classification. Commercial-off-the-shelf (COTS) image processing packages provide an opportunity to identify terrain classification reliability along with the class assignments. However, this opportunity for capturing reliability information is typically neither passed along into the final terrain class map output nor is it stored as a supplementary metadata file. Training sample contingency table scores are used to assign terrain feature class theme, or layer, reliability scores, and distance to class means computed using a Mahalanobis distance technique is used for assignment of within-terrain theme individual pixel reliability scores. For topic two, the value of reliability pixel scores being carried forward into a tactical decision-modeling environment is considered. Reliability from every pixel contributes to a final decision-model product reliability score at the pixel level. An example is provided. In topic three, terrain data layer contribution to model output for HLZ and BIV area tactical decision aids is evaluated on the basis of assessing importance (or weights) for individual or combined terrain layers. Kappa values designating terrain variable contribution are suggested as objective surrogate weights available for use in the HLZ and BIV modeling process.

15. SUBJECT TERMS
digital imagery, terrain data, terrain interpretation, simulation modeling, digital graphic information, terrain modeling, geographic data, military training

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Kevin R. Slocum
Unclassified	Unclassified	Unclassified	SAR	34	19b. TELEPHONE NUMBER (include area code) (703)428-6802

NSN 7540-01-280-5500

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239.18